

Applicants respectfully traverse these rejections.

Applicants teach a system that uses natural language processing (NLP) techniques to automatically assign medical codes (e.g., ICD9 and CPT codes) to diagnoses and procedures performed by physicians and other clinicians described in transcribed reports for billing purposes and the like. The NLP techniques employ a vector space method that utilizes knowledge vectors. The knowledge vectors are manually created by medical and/or linguistic experts. A knowledge vector includes a collection of terms with each term representing one dimension of the vector. Each term in the vector may be weighted based on semantic values. For example, for diagnosis coding, high weight terms may include primary medical conditions and body parts, whereas low weight terms may include adjectives denoting severity and quality.

The system segments a transcribed note and then parses the segments into parse items based on semantic relationships between the concepts. These semantic relationships may be used to determine which individuals are providing information or diagnosis and which are taking actions described in the note. This may be important in clinical environments where the transcribed note may contain the viewpoints of several individuals, for example, the subjective responses of the patient and observations and diagnoses made by medical staff.

These different viewpoints may be treated differently in processing the medical code. Clinical reports may also describe actions taken by different individuals, for example, physicians, nurses, and technicians. The correct identification of the care provider may be necessary for proper billing.

Friedman discloses a natural language extraction system, called MEDLEE, which has been applied to the medical domain. The system may be used to extract text from textual reports, but does not generate standardized medical codes, such as ICD9 or CPT codes. While the system does parse sentences, the parsing is not based on semantic relationships, and hence cannot distinguish between viewpoints and actors represented in a report. One focus of MEDLEE is radiology reports. As noted in the action on page 4, since a radiology report would be unique to a particular individual, recognizing such distinctions may not be important. However, this is not true for, e.g., clinical reports, as discussed above.

The Action states that Friedman fails to expressly disclose the use of knowledge vector processing in generating code. The Action then notes that Friedman does suggest the use of NLP to extract and structure clinical data. Applicants submit that there are many types of NLP techniques, not all of which utilize vector space methods. In fact, the MEDLEE system as implemented

(see <http://cat.cpmc.columbia.edu/medleexml/demo/>) utilizes a linguistic string approach rather than a vector space approach.

Caid discloses a technique for generating context vectors for use in indexing documents. Context vectors differ from knowledge vectors in several significant ways. Context vectors are generated automatically by training a system with a number of documents, e.g., a training corpus. The context vectors generated in this manner are of a fixed length (column 5, lines 16-17) and are weighted based on statistical factors, e.g., frequency and proximity. In contrast, knowledge vectors, as described by Applicants, are generated manually, are of variable length, and are weighted based on semantic rather than statistical factors.

Consider exemplary independent claim 1, which recites in relevant part:

"...(b) applying morphing, parsing and semantic analysis to the segments to generate a normalize file having a standardized form with parse items;

(c) identifying first type matches between parse items of the normalized file and a plurality of standard knowledge vectors..."

Neither Friedman nor Caid, alone or in combination, teaches or suggests, segmenting or parsing a transcribed note based on semantic analysis or identifying type matches based on knowledge

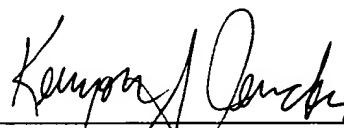
vectors. In fact, Caid teaches away from vector space methods that utilize manually entered vectors, such as the claimed knowledge vectors (column 1, lines 59-64). Accordingly, Applicant submits that the combination of Friedman and Caid is improper.

Applicant asks that all claims be allowed. Enclosed is a \$55 check for the Petition for Extension of Time fee.

Please apply any other charges or credits to Deposit Account No. 06-1050.

Respectfully submitted,

Date: April 22, 2002



Kenyon S. Jenckes
Reg. No. 41,873

Fish & Richardson P.C.
PTO Customer No. 20985
4350 La Jolla Village Drive, Suite 500
San Diego, California 92122
Telephone: (858) 678-5070
Facsimile: (858) 678-5099

10178131.doc